Biases in Food Photo Taking Behavior

Homin Park
National University of Singapore
bighp@nus.edu.sg

Abhinav Ramesh Kashyap
National University of Singapore
e0001169@u.nus.edu

Zhenkai Wang
National University of Singapore
wangzhen@comp.nus.edu.sg

Brian Y. Lim
National University of Singapore
brianlim@comp.nus.edu.sg

Abstract
Recent advances in deep learning have enabled a convenient diet journaling using ubiquitous mobile phones employing automated image recognition systems. These systems are often trained with well-curated images, but users photograph their foods in diverse ways. In this study, we investigate this variation in food photo taking behaviors using a descriptive framework. We collected 5,686 food images from four different types of mobile apps and analyzed them to illustrate biases in terms of the content, context, and photography styles. Our findings contribute to the development and optimization of automated photo-based food tracking algorithms.

Author Keywords
Food Journal; Food Logging; Food Photographs.

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction
Mobile food tracking has been proposed to help manage diet-related chronic diseases, such as diabetes, by increasing users' awareness of their eating behaviors through mobile-based applications [15]. Prior research has investigated the use, challenges, and barriers of mobile apps

(a) Public: nicely photoshopped

(b) Family: communal multi-dish

(c) Professional: has drink with calories

(d) Automated: tightly-cropped image

Figure 1: Representative images found in different app types.

for food logging [3, 4]. However, to date, there has been no study conducted to comprehend how people capture food photos and its impact on diet assessments. Given that photo-based food classifiers are often trained with carefully selected and cleaned images to perform accurately, the variation in actual food photographs will introduce bias and degrade recognition accuracy.

To investigate biases, we analyzed 5,686 of images from four types of mobile applications: automated image recognition (Nibble [12] and FoodLg); photo logging with professional dietitian support (Glycoleap [6]); food journaling as a family (TableChat); and image-based social media (Instagram with #fooddiary and #foodjournal tags [3]). Specifically, these apps vary by social context of who assesses the food log information: no one (auto), professionals (pro), family, and public. We describe each food image with 19 attributes, grouped into three categories (content, context, and style (see Table 1)).

Influencing Factors on Food Photo-taking

Social cognitive theory suggests that behavioral interventions via cognitive mediators (e.g., self-efficacy, outcome expectation, and user intention) could influence our health-related behaviors [1]. We draw on two theoretical frameworks, namely, social influence and user expectation to reason why there are biases in food photo taking behaviors.

Social Influence

Glanz et al. [5] summarized a number of behavioral science theories implemented in public health interventions to promote healthier behavior changes. Among various theories addressed, social cognitive theory [1] and social ecological model [17] best articulates how user behaviors are shaped by the social environment (such as individual, interpersonal, organizational, community, and public). This has also been proven by various recent studies, including one conducted by Kato-Lin et al., which illustrates the influence of peer types (peers with existing ties and one without) on the term of engagements and changes in eating behaviors [9]. In addition, dietitians are a credible source for dietary assessment, which can lead to increased user engagement because instructions and feedbacks from authoritative professionals can have greater impact on behavior changes [2, 18].

User Expectation

User expectations on how automatic food recognition systems would perform can also influence how photos are captured. Users form expectations about an interactive technology even before they use it and this affects the way they interact with it [13]. Similarly, Lee and See [11] found that the level of user reliance on automated system depends on how well user's trust matches the true capabilities of automation. Users are likely to go through multiple calibration phases to close the gap between the level of trust and the capabilities of automation [10, 14]. This concept of trust follows a social learning theory where expectations are determined by previous experiences that are perceived to be similar [16]. In the case of automatic food recognition logging, some users may have limited initial trust and only test the system simple photographs (e.g., one item on a plate). Thus, user expectation can further inform us about the different photographing behaviours found in different photo-based food logging applications.

Analysis and Results

Five researchers labeled the images across 19 attributes. For each attribute, we performed a one-way ANOVA with app type as the single factor with four levels (Public, Family, Pro, and Auto). We applied Bonferroni correction due to the large numbers of metrics being compared. Table 1 summa-

Table 1: Statistical results and occurrence rates of each descriptive measure across different app types. Superscripts A, B, C, and D indicate significantly different groups found for the measure by a Tukey HSD test ($\alpha = 0.0027$).

		p	**Public**	**Family**	**Pro**	**Auto**
Content	Is Food	<.0001	75%^{C}	98%^{A}	89%^{B}	98%^{A}
	# Dishes	<.0001	1.6^{B}	2.2^{A}	1.5^{B}	1.2^{C}
	# Parts/Dish	<.0001	1.7^{B}	1.7^{AB}	1.8^{A}	1.4^{C}
	Has Drink	<.0001	17%^{B}	20%^{ABC}	23%^{A}	11%^{C}
	Grocery	<.0001	11%^{A}	5%^{AB}	1%^{B}	1%^{B}
	Packaged	<.0001	19%^{A}	8%^{B}	8%^{B}	4%^{B}
Context	Is Healthy	<.0001	80%^{B}	74%^{C}	83%^{A}	54%^{D}
	Multi-course	<.0028	0.8%^{A}	1.3%^{A}	0.1%^{A}	0.2%^{A}
	Communal	<.0001	2.5%^{B}	18%^{A}	1.9%^{B}	4.3%^{B}
	Has Other's Food	<.0001	2.4%^{B}	15%^{A}	2.9%^{B}	2.8%^{B}
	Partially Eaten	<.0001	3.1%^{B}	7.1%^{AB}	5.1%^{B}	12%^{A}
	Handheld	<.0001	6.5%^{A}	2.2%^{AB}	2.9%^{B}	3.4%^{AB}
	Has Faces	<.0001	1.3%^{A}	0%^{AB}	0.2%^{B}	0.4%^{AB}
Style	Top-down	<.0001	65%^{A}	70%^{A}	65%^{A}	42%^{B}
	Multi-shot	<.0001	0.3%^{B}	0.9%^{A}	0.6%^{B}	5.6%^{B}
	Tight-cropped	<.0001	6%^{B}	3%^{B}	7%^{AB}	11%^{A}
	Collage	<.0001	3%^{A}	0%^{AB}	2%^{B}	0%^{B}
	Photoshopped	<.0001	49%^{A}	0%^{B}	0%^{B}	2%^{B}
	Poor Lighting	<.0001	5%^{B}	8%^{AB}	6%^{B}	12%^{A}

rizes the results of the Tukey HSD tests and illustrates the prevalence of each attribute.

Our results reveal that images found from the public app are frequently photoshopped and have least amount of partially eaten foods compared to the others. Whereas, images found in the family app tends to include communal and shared tables increasing the average number of dishes and parts per dish. For the professional app, we found that most users frequently post their drinks. Lastly, automated apps have most number of tight-cropped and poor lighted images. These characteristics are illustrated in Figure 1.

Design Implications

Findings from our analysis results shown in Table 1 can be utilized to improve and address following three aspects of automated photo-based food logging apps. By combining the three, we could mitigate issues raised from the biases in content, context, and style of the photos.

Drink Classifier and Dataset: Content

We find that a significant proportion (20%) of the photos have drinks in them. The current food recognition technologies that count calories do not include drinks and would lead to inaccurate mobile food logging. Drinks are often much more difficult than foods to be categorized because of its limited visual cue. Unlike foods where some ingredients are visible, drinks conform to the shape of its container thus losing significant amount of important information. We suggest that future mobile food logging technologies consider developing dedicated drink classifier and curating a dataset specifically targeting drinks only.

Guidelines for Food Photo Taking: Context and Style

Our results indicate that users take food photos in diverse ways. There are some apps that help the user to frame photos to increase its attractiveness [8], but here we suggest automated food logging app designers to consider providing comprehensive guidelines to assist users who want to receive detailed nutritional analysis to take proper food photos. Interventions could inform the user if the lighting is poor, or assist the user to not crop the images excessively. By following the guidelines, users will be able to ensure good assessment accuracy.

De-biasing the captured photos: Style

Unlike Instagram, Google, Facebook, and other major companies, most developers do not have the dataset that is captured from the field and is big enough to ensure accurate food recognition and assessment results. In fact, most development teams utilize publicly available open image datasets to train their initial classification models, and gradually improve their system as more images are collected by their users. To improve the accuracy of the initial classifier, we suggest developers to pre-process the captured image to look more like the training dataset. Latest generative adversarial network approaches [7] can be employed to handle such a task.

Conclusion

Our work contributes to informing the development and optimization of automated photo-based food logging algorithms by illustrating the bias in food photo taking under different social contexts. With the descriptive framework that we have developed, designers and developers can identify differences in food photos across apps, and develop algorithms to handle drink classifications, guide food photo taking, and de-bias the photos towards curated training image datasets. In future, we aim to develop more accurate food image recognition models to improve accuracy in the field.

REFERENCES

1. Albert Bandura. 1998. Health promotion from the perspective of social cognitive theory. *Psychology and health* 13, 4 (1998), 623–649.
2. Wendy Brouwer, Willemieke Kroeze, Rik Crutzen, Jascha de Nooijer, Nanne K de Vries, Johannes Brug, and Anke Oenema. 2011. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. *Journal of medical Internet research* 13, 1 (2011).
3. Chia-Fang Chung, Elena Agapie, Jessica Schroeder, Sonali Mishra, James Fogarty, and Sean A Munson. 2017. When Personal Tracking Becomes Social: Examining the Use of Instagram for Healthy Eating. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 1674–1687.
4. Felicia Cordeiro, Daniel A Epstein, Edison Thomaz, Elizabeth Bales, Arvind K Jagannathan, Gregory D Abowd, and James Fogarty. 2015. Barriers and negative nudges: Exploring challenges in food journaling. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 1159–1162.
5. Karen Glanz and Donald B Bishop. 2010. The role of behavioral science theory in development and implementation of public health interventions. *Annual review of public health* 31 (2010), 399–418.
6. GlycoLeap. 2016. diabetes management app by Holmusk. https://glycoleap.com/. Retrieved 21 September 20, 2016. (2016).
7. Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems*. 2672–2680.
8. Takao Kakimori, Makoto Okabe, Keiji Yanai, and Rikio Onai. 2015. A system to support the amateurs to take a delicious-looking picture of foods. In *SIGGRAPH Asia 2015 Mobile Graphics and Interactive Applications*. ACM, 28.
9. Yi-Chin Kato-Lin, Vibhanshu Abhishek, Julie S Downs, and Rema Padman. 2016. Food for Thought: The Impact of m-Health Enabled Interventions on Eating Behavior. (2016).

10. John D Lee and Neville Moray. 1994. Trust, self-confidence, and operators' adaptation to automation. *International journal of human-computer studies* 40, 1 (1994), 153–184.

11. John D Lee and Katrina A See. 2004. Trust in automation: Designing for appropriate reliance. *Human factors* 46, 1 (2004), 50–80.

12. Brian Y Lim, Xinni Chng, and Shengdong Zhao. 2017. Trade-off between Automation and Accuracy in Mobile Photo Recognition Food Logging. In *Proceedings of the Fifth International Symposium of Chinese CHI*. ACM, 53–59.

13. Jaroslav Michalco, Jakob Grue Simonsen, and Kasper Hornbæk. 2015. An exploration of the relation between expectations and user experience. *International Journal of Human-Computer Interaction* 31, 9 (2015), 603–617.

14. Bonnie M Muir. 1987. Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Studies* 27, 5-6 (1987), 527–539.

15. Jessica A Pater, Oliver L Haimson, Nazanin Andalibi, and Elizabeth D Mynatt. 2016. âĂIJHunger Hurts but Starving WorksâĂİ: Characterizing the presentation of eating disorders online. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 1185–1200.

16. Julian B Rotter. 1971. Generalized expectancies for interpersonal trust. *American psychologist* 26, 5 (1971), 443.

17. James F Sallis, Neville Owen, and E Fisher. 2015. Ecological models of health behavior. *Health behavior: Theory, research, and practice* 5 (2015), 43–64.

18. Janine L Wright, Jillian L Sherriff, Satvinder S Dhaliwal, and John CL Mamo. 2011. Tailored, iterative, printed dietary feedback is as effective as group education in improving dietary behaviours: results from a randomised control trial in middle-aged adults with cardiovascular risk factors. *International Journal of Behavioral Nutrition and Physical Activity* 8, 1 (2011), 43.